TENTAMENSKOD:________________

INSTITUTIONEN FÖR
NATURVETENSKAP OCH TEKNIK

Kursens namn/kurskod	**Numeriska Metoder för civilingenjörer** **DT508G**
Examinationsmomentets namn/provkod	**Teori, 3,5 högskolepoäng (Provkod: A001)**
Datum	2020-08-20
Tid	Kl. 08:15 – 13:15

Tillåtna hjälpmedel	Skrivmateriel, formel blad och miniräknare med raderat minne, kursbok, egna anteckningar från föreläsning och datorlaboration.
Instruktion	Läs igenom alla frågor noga. Börja varje fråga på ett nytt svarsblad. Skriv bara på ena sidan av svarsbladet. Skriv tentamenskoden på varje svarsblad. Skriv läsligt! Det räcker att ange fem decimaler. Glöm inte att scanna lösningar och ladda up det till WISEflow
Viktigt att tänka på	Motivera väl, redovisa tydligt alla väsentliga steg, rita tydliga figurer och svara med rätt enhet. Lämna in i uppgiftsordning.
Ansvarig/-a lärare (ev. telefonnummer)	Danny Thonig (Mobil: 0727010037)
Totalt antal poäng	40
Betyg (ev. ECTS)	Skrivningens maxpoäng är 40. För betyg 3/4/5 räcker det med samt 20/28/38 poäng totalt. Detaljerna framgår av separat dokument publicerat på Blackboard.
Tentamensresultat	Resultatet meddelas på Studentforum inom 15 arbetsdagar efter tentadagen.
Övrigt	Lärare är inte på plats. Varsågod och ringa om du har frågar.

Lycka till!

1. Decide whether the following statements are true or false. Explain the reason for your answer:

a) Doolittle's method for LU factorization requires 1's in the diagonal of the U matrix, where Crout's method requires 1's in the diagonal of the L matrix. [1p]

b) Let f be a continuously differentiable function on $[a, b]$ and $x_0, x_1, \ldots, x_n \in [a, b]$ are distinct, i.e., $x_i \neq x_j, i \neq j$. Then the Hermite polynomial that interpolates $(x_0, f(x_0)), \ldots, (x_n, f(x_n))$ has the degree $2n + 1$. [1p]

c) The Newton method diverges for the function $f(x) = x^a$, when $0 < a < \frac{1}{2}$. [1p]

d) The roots of the Wilkinson polynomial of the order n (n Is large and e.g. 10) can be calculated exact by numerical methods. [1p]

e) The machine error for single precession numbers is smaller than 10^{-10}. [1p]

f) A root r of $f(r) = 0$ is called simple if the multiplicity is zero. [1p]

g) Let an iterative method for nonlinear equations generate a sequence $\{x_n\}$ that converges to the limit r and has the error $e_n = |x_n - r|$, than the method's convergence order p is defined by $0 < \frac{pe_{n+1}}{e_n} < \infty$. [1p]

h) The second derivative $y''(x)$ can be approximated by $y'' = \frac{y_{n+1} - y_{n-1}}{h}$, where h is the numerical step-width between two point x_n and x_{n+1}. [1p]

i) The maximum possible error magnification factor for solving $Ax = b$ is related to the product of forward and backward error. [1p]

j) A symmetric matrix A is positive definite if and only if Gauss elimination without row interchange can be performed on the linear system $Ax = b$ with all pivot elements positive. [1p]

Suggested Solution:

a) False, in Doolittle L have 1's in diagonal and in Crout's method, U has 1's in diagonal
b) True, since Hermite interpolation involves also derivatives
c) True, see example in the book.
d) False, since e.g. the constant a_0 in the polynomial is already $n!$ and calculating root would be not accurate due to rounding errors.
e) False, $2^{-23} = 1 \cdot 10^{-6}$
f) False, it is called simple if the multiplicity is one.
g) False, $0 < \frac{e_{i+1}}{e_i^p} < \infty$
h) False, it is the definition for the first derivative
i) False, it is the ration between forward and backward error
j) True, see theorem in the book

2. Let function $f \in C^{n+1}[a, b]$, $|f^{(n+1)}(c(x))| \leq M$ and $e_{n+1} = |f(x) - P_{n+1}(x)|$ (see Formula sheet) be the error of its best approximation by a polynomial $P_{n+1}(x)$ of degree $n + 1$. Show that the accuracy of the best polynomial approximation improves rapidly as the size of the interval $[a, b]$ shrinks, i.e., show that

$$e_{n+1} \leq \frac{2M}{(n+1)!}\left(\frac{b-a}{4}\right)^{n+1}.$$

Use the Chebyshev nodes to construct a polynomial approximation of f. [3p]

Suggested Solution:

The error term of interpolation by a polynomial if degree $n + 1$ is

$$\left|f(x) - P_{n+1}(x)\right| = \left|\frac{(x-x_1)(x-x_2)\ldots(x-x_{n+1})}{(n+1)!} f^{(n+1)}(c(x))\right| \leq \max_c \frac{|f^{(n+1)}(c(x))|}{(n+1)!}\Pi_{i=1}^{n+1}(x-x_i)$$

Where x_i are the $n + 1$ nodes at which the values of f are given. We use Chebyshev nodes for the x_i

$$x_i = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)\cos(\frac{\pi}{2}\frac{2i-1}{2n+1})$$

We change the variable from x to y, where $y \in [-1,1]$ and

$$x = x(y) = \frac{1}{2}(b+a) + \frac{1}{2}(b-a)y$$

We see directly that the new nodes y_i are $\cos(\frac{\pi}{2}\frac{2i-1}{2n+1})$ and thus

$$\left|f(x) - P_{n+1}(x)\right| \leq \max_c \frac{|f^{(n+1)}(c(x(y)))|}{(n+1)!}\Pi_{i=1}^{n+1}\frac{b-a}{2}(y - \cos(\frac{\pi}{2}\frac{2i-1}{2n+1}))$$

The last product is known as Chebyshev polynomial T_{n+1}

$$\Pi_{i=1}^{n+1}(y - \cos(\frac{\pi}{2}\frac{2i-1}{2n+1})) = \frac{T_{n+1}(y)}{2^n}.$$

We arrive at

$$\left|f(x) - P_{n+1}(x(y))\right| \leq \max_c \frac{|f^{(n+1)}(c(x(y)))|}{(n+1)!}\left(\frac{b-a}{2}\right)^{n+1}\frac{|T_{n+1}(y)|}{2^n} \leq \frac{2M}{(n+1)!}\left(\frac{b-a}{4}\right)^{n+1}$$

If the size of the interval $b - a$ shrinks, the error shrinks by $(b-a)^{n+1}$.

3. Arctic sea ice reaches its minimum each September. September Arctic sea ice is now declining at a rate of 12.85 percent per decade. The table shows the average monthly Arctic sea ice extent each September since 1990, derived from satellite observations. To obtain a functional trend, the data should be approximated by least squares polynomials of degrees 2. Data is obtained from from https://climate.nasa.gov/vital-signs/arctic-sea-ice/

   a) Formulate the problem into a system of equations $Ax = b$. Why it is impossible to

i	Year	Arctic sea ice (in million sq km)
1	1990	6.14
2	1995	6.08
3	2000	6.25
4	2005	5.50
5	2010	4.87
6	2015	4.62

   determine an exact solution? What is the least square solution? [3p]

   b) Solve the normal equation for the problem. [4p]

Suggested Solution:

a)
A polynomial of order 2 is
$P(x) = a_2x^2 + a_1x + a_0$

Assuming a least square fit, using the years probably overloads the problem. This, we can use the first column of integers.

Plug in the integers in the polynomial as x value and replacing $P(x)$ by the values of the gold prices, we obtain

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \\ 25 & 5 & 1 \\ 36 & 6 & 1 \end{pmatrix} x = \begin{pmatrix} a_2 \\ a_1 \\ a_0 \end{pmatrix} \text{ and } b = \begin{pmatrix} 6.14 \\ 6.08 \\ 6.25 \\ 5.50 \\ 4.87 \\ 4.62 \end{pmatrix}. \text{ [1p]}$$

There are more equations (6) than unknowns (3) and no solution exists [1p]. However, least square problem can be defined, that minimizes the residual $r = ||b - A\bar{x}||$. For more information on the least square method, see book. [1p]

For the least square problem, the following equation has to be solved
$A^T A\bar{x} = A^T b$

b)

Thus,

$$A^T A = \begin{pmatrix} 1 & 4 & 9 & 16 & 25 & 36 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \\ 25 & 5 & 1 \\ 36 & 6 & 1 \end{pmatrix} = \begin{pmatrix} 2275 & 441 & 91 \\ 441 & 91 & 21 \\ 91 & 21 & 6 \end{pmatrix}$$

$$A^T b = \begin{pmatrix} 1 & 4 & 9 & 16 & 25 & 36 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 6.14 \\ 6.08 \\ 6.25 \\ 5.50 \\ 4.87 \\ 4.62 \end{pmatrix} = \begin{pmatrix} 462.78 \\ 111.12 \\ 33.46 \end{pmatrix}$$

The solution of the system of linear equations $A^T A \bar{x} = A^T b$ is

$$\bar{x} = \begin{pmatrix} -0.0741 \\ 0.1765 \\ 6.0830 \end{pmatrix}$$

4. Let

$$A = \begin{bmatrix} 2 & -9 & 0 & 4 \\ 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ 2 & 1 & 0 & 0 \end{bmatrix}$$

be coefficient matrix of systems of linear equations $Ax = b$.

a) Describe one iterative and one non-iterative method to solve the system of linear equations? Compare both problems in terms of computational costs and efficiency. [2p]

b) Perform LU factorisation of $A = LU$ in the Crout's method. [2p]

c) Apply the obtained LU factorisation of A to calculate the inverse of the matrix, A^{-1}. [2p]

d) Compute the condition number of the matrix A with respect to the max norm. How much larger can the solution's relative error $\frac{||\delta x||_\infty}{||x||_\infty}$ be compared to the data error $\frac{||\delta b||_\infty}{||b||_\infty}$ when solving $Ax = b$. [2p]

e) Suppose that in the system above $b = (0\ \beta\ 0\ 0)^T$, where the zero elements are exact and contain no error, while β is correctly rounded. How much larger us the solution's maximum relative error now, still measured in the max norm? [2p]

Suggested Solution:

a)

E.g. Gauss-Seidel as an iterative method and PA=LU decomposition as a non-iterative method. For more details, see book.

b)

Crout's method need to solve

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & u_{13} & u_{14} \\ 0 & 1 & u_{23} & u_{24} \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$l_{11} = a_{11} = 2, l_{11}u_{12} = a_{12} \rightarrow u_{12} = \frac{-9}{2}, l_{11}u_{13} = a_{13} \rightarrow u_{13} = 0, l_{11}u_{14} = a_{14} \rightarrow u_{14} = 2$

$l_{21} = a_{21} = 1, l_{21}u_{12} + l_{22} = a_{22} \rightarrow l_{22} = \frac{11}{2}$ and so forth. Thus

$$L = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & \frac{11}{2} & 0 & 0 \\ 0 & -1 & \frac{2}{11} & 0 \\ 2 & 10 & -\frac{20}{11} & 6 \end{bmatrix}, U = \begin{bmatrix} 1 & \frac{-9}{2} & 0 & 2 \\ 0 & 1 & \frac{2}{11} & -\frac{2}{11} \\ 0 & 0 & 1 & \frac{9}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

c)

For matrix inversion of the 4×4 matrix, we have to solve the tableau

$$\begin{bmatrix} 2 & -9 & 0 & 4 & | & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & | & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & | & 0 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & | & 0 & 0 & 0 & 1 \end{bmatrix}$$

This implies to solve four systems of linear equations $Ax = b_i\,,\ i = 1,\ldots,4$. For simplicity we can use the LU factorization as $Ax = b \rightarrow Lc = b, Ux = c$. Thus we have in total eight trivial systems of equations with triangular matrices L and U.
The final result is

$$A^{-1} = \frac{1}{12} \begin{bmatrix} 1 & 0 & -4 & 5 \\ -2 & 0 & 8 & 2 \\ 3 & 12 & -24 & -9 \\ -2 & 0 & 20 & 2 \end{bmatrix}$$

d)

The condition number is $cond(A) = ||A||_\infty ||A^{-1}||_\infty = 15 \cdot \frac{48}{12} = 60$, so the solution's relative error can be at most 60 times larger.

e)

If $b = (0\ \beta\ 0\ 0)^T$, then $x = (0\ 0\ \beta\ 0)^T$ (cf. the structure of the inverse matrix A^{-1}). Therefore b and x have the same relative error, so the factor is exactly 1.

5. The fundamental theorem of interpolation says that each function can be approximated by a polynomial, such as the Lagrange polynomial. Based on this, solve the following part problems
   a) Derive the Three-Point Formula for the second derivative $f''(x)$ using equally spaced point $x_0 = x_1 - h$, x_1, and $x_2 = x_1 + h$ and the Lagrange polynomial. You can check the results with the one on the formula sheet. [4p]
   b) Calculate the error of numerical rule found above. What is the order of approximation? [3p]

Suggested Solution:

a)

From

$$f(x) = \sum_{k=1}^{n} f(x_k)L_k(x) + \frac{(x-x_1)(x-x_2)\ldots(x-x_n)}{n!} f^{(n)}(c(x))$$

$$L_k(x) = \frac{(x-x_1)(x-x_2)\ldots\overline{(x-x_k)}\ldots(x-x_n)}{(x_k-x_1)(x_k-x_2)\ldots\overline{(x_k-x_k)}\ldots(x_k-x_n)}.$$

We obtain for the three points

$$f(x) \approx y_0\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} + y_1\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + y_2\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}.$$

Here $y_k = f(x_k)$. The second derivative of this function is

$$f''(x) \approx \frac{2y_0}{(x_0-x_1)(x_0-x_2)} + \frac{2y_1}{(x_1-x_0)(x_1-x_2)} + \frac{2y_2}{(x_2-x_0)(x_2-x_1)}$$

Lets call for less confusion $x_1 = y$ and, thus, $x_0 = y - h$, $x_1 = y$, and $x_2 = y + h$ with the numerical step width h. Consequently,

$$f''(y) \approx \frac{2f(y-h)}{(-h)(-2h)} + \frac{2f(y)}{(h)(h)} + \frac{2f(y+h)}{(2h)(h)}$$

And, finally

$$f''(y) \approx \frac{f(y-h) - 2f(y) + f(y+h)}{h^2}.$$

b) The error is

$$e(y) = \frac{1}{6}\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}\Pi_{k=0}^{2}(x-x_k)\right)|_{x=y} f^{(3)}(c(x))$$

Assuming $\dfrac{d^2}{dx^2}f^{(3)}(c(y)) = 0$ and, thus,

$$e(y) = 0$$

6. Compute numerically the integral

$$\int_0^2 \frac{\sin x}{1+x}\mathrm{d}x$$

to an accuracy of at least 10^{-5} and show by estimating the error that you have achieved this accuracy. [3p]

Suggested Solution:

We use Romberg's method, and start with step size $h = \dfrac{1}{2}$ and then use $h = \dfrac{1}{2}$ and $h = \dfrac{1}{2}$. We then get the Trapezoidal sums

$T(\dfrac{1}{2}) = 0.64545004$, $T(\dfrac{1}{4}) = 0.66447878$, $T(\dfrac{1}{8}) = 0.66929831$

We note that the differences are regular and have a quotient of 4, which is expected as the integrand is an analytic function. Extrapolating to eliminate the $O(h^2)$ error produces

$R(\dfrac{1}{4}) = 0.67082170$, $R(\dfrac{1}{8}) = 0.67090482$

Here we extrapolate to eliminate the $O(h^4)$ error; the correction term is $5.5 \cdot 10^{-6}$ and the result is $S(\dfrac{1}{8}) = 0.67091036$. Since the correction term was smaller than the requested accuracy, we accept the answer 0.67091.